

### C. Remarks

#### Status of Pending Claims

Claims 1 – 8 and 26 – 36 are currently pending in the present application.

Claim 30 has been amended to correct a clerical error by removing the indicated phrase at line 7 of the claim that is duplicative of the phrase at line 14. Claim 30 has also been amended to correct a minor failure of ambiguous antecedence basis at line 10. Claims 33 and 34 have also been amended to maintain correct antecedent basis relation with Claim 30.

#### Restriction Requirement

The election of the Group I set of claims (Claims 1 – 8 and 26 – 36) is affirmed. Accordingly, the Group II claims (Claims 9 – 25 and 37 – 43) have been withdrawn from further consideration in the present application.

#### Claim Objections

At the direction of the Examiner, Claim 33 has been amended to insert a comma at line 5 after the word 'mesh.' Withdrawal of the object is respectfully requested.

#### Information Disclosure Statement

The patent and literature documents identified in the enclosed IDS are believed to be no more relevant than the documents identified by the Examiner. They are submitted for completeness and as background. Electronic copies of the literature documents can be provided at the request of the Examiner.

#### Rejections under 35 U.S.C. §102:

Claims 1, 2 and 30 – 36 were rejected under 35 U.S.C. §102 as anticipated by Kaufman (US Patent 6,240,408).

In order to establish a rejection under 35 U.S.C. §102, all elements of a claim must be identically found in a prior art reference. See, M.P.E.P. §706.02 (For anticipation under 35 U.S.C. 102, the reference must teach every aspect of the claimed invention either explicitly or impliedly. Any feature not directly taught must be inherently present) (emphasis added); M.P.E.P. §2112 (In relying upon the theory of inherency, the Examiner must provide a basis in fact and/or technical reasoning to reasonably support the determination that the allegedly inherent characteristic necessarily flows from the teachings of the applied prior art. Ex parte Levy, 17 USPQ2d 1461, 1464 (Bd. Pat. App. & Inter. 1990) (emphasis in original); M.P.E.P. §2131.

The essential nature of anticipatory identity requires that the function of the elements and their interconnections not just be colorably similar, but identical in all aspects (emphasis added). See, Richardson v. Suzuki Motor Co., 868 F.2d 1226, 1236, 9 USPQ2d 1913, 1920 (Fed. Cir. 1989) (The identical invention must be shown [by the reference] in as complete detail as is contained in the . . . claim). Clearly, a prior art reference that discloses a collection of elements that are assembled differently and that function collectively in a different or incomplete way compared to the claimed invention is not an anticipating reference.

#### Claim 1:

Kaufman teaches an authoritative database that contains “documents selected to be reference materials of high reliability . . . includ[ing] encyclopedias, almanacs, dictionaries and other reference materials . . . subject to professional editorial review.” (Col 4, ll 16-21.) These documents are clearly the full text documents themselves. That is, they contain all of the sentences, otherwise undifferentiated, that exist within the included reference material documents.

The Kaufman database is “authoritative” only in the sense that the selection of documents contained is considered “more reliable reference materials” as compared to a completely uncontrolled public database. The Kaufman databases are parsed in parallel in response to dynamically provided user queries. The Kaufman authoritative database does not substitute for or reference the public database, but rather only serves as a supplementary source of potential query matches.

To evaluate an arbitrary query text, as dynamically presented by a user, Kaufman teaches that the documents of the authoritative (and public) database are retrieved and parsed to find word matches. Various values based on word-in-document frequency and word-in-sentence frequency are computed. Of significance, Kaufman explicitly teaches that this analysis is performed dynamically, as is required given that the specific user asserted query text is not otherwise known in advance. Sentence boundaries are taught as found dynamically (col 6, ll 1-8). The various frequency values are computed dynamically (col 6, ll 28 - 36). The reference makes no mention of any information stored in the authoritative database other than the simple full text of the documents themselves.

Finally, Kaufman teaches that the results of a query are ordered based on the values found and presented in a simple, linear list (col 12, ll 46 - 48).

In contrast, the preamble of Claim 1 provides for a computer system that enables research to be conducted specifically against an “authoritatively organized document collection.” As defined in the present specification, such an authoritative document collection employs

. . . citations to internally organize and substantiate the information represented by the collection. Such authoritative document collections, including as exemplary the various scientific and legal document collections . . . (¶48)

Not all document collections are “authoritative document collections.” Notably, much of the reference material specifically identified by Kaufman for inclusion in its authoritative database are not authoritative document collections, namely encyclopedias, almanacs, and dictionaries. For at least this reason, the authoritative database of Kaufman is not identical to and therefore cannot be considered to anticipate the database as set forth in Claim 1.

Claim 1 further requires that the claimed database store

. . . first data identifying a set of authoritative statements present within the documents of said predetermined authoritative document collection

As defined in the present specification (¶48), an “authoritative statement” is defined as a portion of text that includes authoritative “assertions and citations.” Such authoritative assertions are defined by the present specification as

. . . a statement made to impliedly establish a concept or contention as fact, typically supported by citation reference to a preexisting basis or line of reasoning, typically associated with a prior or precedential authoritative assertion, or statement of convention, such as a statute or definition. (¶61).

Thus, the database itself must store “first data” that references sets of authoritative statements, each of which includes at least one assertion and one supporting citation.

Contrary to the implied assertion made in the Action, the claimed first data is not and cannot be simply the otherwise undifferentiated collection of sentences that exist in the documents of the document collection.

Rather, Claim 1, through the defined meaning of the claim terms used, specifies and requires a particular relationship between the first data stored by the database and a specific authoritative document collection. The claimed first data is required to identify the differentiated authoritative assertions that exist within and among the sentences that make up the documents of an authoritative document collection. In that sense, the claimed first data is itself not part of the authoritative document collection. Claim 1 requires database storage of this first data.

The Kaufman authoritative database stores no counterpart to the claimed first data. The Kaufman authoritative database only stores the full text of the reference material documents. Kaufman does not teach recognition or differential treatment of authoritative assertions. Kaufman does not teach the pre-identification of any authoritative assertions as first data and the storage of that first data to a database.

Since Kaufman does not identically teach the database storage of first data identifying authoritative assertions, Kaufman cannot be properly considered to anticipate Claim 1.

The database of Claim 1 further stores

. . . second data specifying the locations of the authoritative assertions of said set of authoritative assertions within the documents of said predetermined authoritative document collection

The second data explicitly functions to locate authoritative assertions among the various sentences within the documents of an authoritative document collection. This location data is pre-determined and pre-stored to the claimed database. Kaufman, in contrast, only teaches the dynamic quantization of sentences – all sentences completely undifferentiated by any categorization – upon parsing in response to the presentation of a specific query text (col 6, ll 1-8).

For at least this reason, Kaufman does not identically teach Claim 1 and, therefore, cannot be properly considered as anticipating the claim.

The database of Claim 1 contains the further requirement of storing

. . . third data specifying correlated associations between the authoritative assertions of said set of authoritative assertions within the documents of said predetermined authoritative document collection

Kaufman does not teach that its authoritative database stores correlated associations between selected parts of any documents relative to the documents that contain those parts. Kaufman only teaches the dynamic generation of associations by a data analyzer in response to a particular query. The product of the Kaufman data analyzer is simply discarded once used to order the list of search results.

Since Kaufman does not identically teach the storage the claimed third data, Kaufman cannot be properly considered to anticipate Claim 1.

Finally, Claim 1 requires a processor that, based on the information as stored in the claimed database, operates

. . . to generate a mesh representational view of the correlated associations between the authoritative assertions of said set of authoritative assertions

Kaufman plainly – and only – teaches presenting a “list” of documents representing the search results (col 12, ll 46 – 48). In contrast, Claim 1 is explicit in requiring the generation of a “mesh” representational view of the “correlated associations.” As plainly understood in the art, a list is a one-dimensional data structure while a mesh is a two-or-more dimensional data structure as described and consistently used in the present specification. Moreover, the claimed mesh is generated by the processor coupled to the database for presentation to a user, as is well evident from the full term “mesh representational view.”

The implication in the Action that the drawings presented in the Kaufman patent document itself are “representational views” fails for many reasons, not the least being that the patent drawings and the relations they present are never generated as “representational views” by the Kaufman system itself or presented to any user of the Kaufman system.

For this as well as the several reasons discussed above, the system taught by Kaufman is not identical to the computer system set forth in Claim 1. Applicant therefore respectfully asserts that Claim 1 is not anticipated by Kaufman. Reconsideration of the rejection of Claim 1 is respectfully requested.

#### Claim 2:

Claim 2 further qualifies the third data as “defin[ing] relative distance weighted, directional associations between the authoritative assertions.”

As dependent on Claim 1, Claim 2 requires the third data to be stored in the database. Further, the associations defined by the third data are required to be specific to authoritative assertions within the documents of an authoritative document collection.

The Action references nearly five full columns of the reference in support of the rejection of this claim. While all generally relevant – the portions do speak to the computation of various metrics – unfortunately, no specific relevance to the particular requirements of the claim can be reasonably discerned. Clarification for this claim, and for

Claims 31 – 36 against which exactly the same portions of the reference are cited, is respectfully requested.

The metrics computed by Kaufman are ultimately combined to generate similarity rankings. These metrics include various statistical values, such as word frequency correlated by document and sentence. According to Kaufman, these metrics are dynamically generated, combined to determine document rankings, and evidently discarded. In computing the rankings, Kaufman mentions assigning additional weight to sentence associated metrics based on the position of the sentence within a document – possibly weighting a sentence more if the position is closer to the document beginning or end. Again, these supplemental weightings are evidently discarded on computation of the document ranking.

Notably, Kaufman does not teach computation of any weightings that embody consideration of both the relative distance and direction between the matched words or sentences of a document. Note that the sentence position-dependent calculation of Equation 5 (col 11, ln 40) sums weights variable relative only to the end of the document; the mutual position of any two matches is never considered. The direction between any two matches is never considered in Kaufman. Furthermore, Kaufman certainly does not teach the storing of any such computed values to a database.

Applicant respectfully asserts that Kaufman cannot anticipate Claim 2. Reconsideration of the rejection of Claim 2 is respectfully requested.

#### Claim 30:

For reasons similar to Claim 1, Claim 30 is also not anticipated by Kaufman. Specifically, Claim 1 requires a processor to generate and store “first” and “second” data to a database. This data is specified as related to the occurrence of “authoritative assertions” within a document collection. For clarity, the claim has been amended to explicitly state that the document collection is an “authoritative document collection,” though such is implicit given that the claim as originally presented requires identification of authoritative assertions; as defined in the present specification, authoritative assertions exist only in authoritative document collections.

The first data is explicitly specified by the claim as both “identifying said set of authoritative assertions” and “identifying the locations of said set of authoritative assertions within the documents of a predetermined document collection.”

Kaufman plainly does not teach the identification of authoritative assertions and certainly does not teach the storage of such identifications to any database.

Claim 30 also specifies the second data as

. . . containing a weighted correlation of the mutual relative occurrence of the authoritative assertions of said set of authoritative assertions within the documents of said predetermined document collection

Again, Kaufman does not teach recognition of authoritative assertions, and clearly does not teach the creation and storage of correlated weighted associations between authoritative assertions. To meet the requirements of the claim, the weighted correlations must be determinable from the documents of an authoritative document collection – not dependent on the content of a user provided query. Kaufman, however, teaches computation of weights and similarity rankings dynamically upon presentation of a query text and then entirely dependent on the content of the query.

Kaufman thus fails to identically teach the generation of the claimed first and second data and, further, the storage of such data as generated. Applicant therefore respectfully asserts that Kaufman does not anticipate the computer system set forth in Claim 30. Reconsideration of the rejection of Claim 30 is respectfully requested.

#### Claim 31:

Claim 31 specifies that the second data “contains weighted correlations representing semantic similarity of the authoritative assertions.” Applicant could not find the word “semantic,” or words of similar conceptual reference, anywhere in the text of the reference. Clarification of the basis for rejection is requested.

In any event, Applicant respectfully asserts that Kaufman does not teach any semantic analysis, let alone semantic analysis specifically in relation to authoritative assertions. As



discussed in relation to Claims 1 and 2, Kaufman does not even recognize the existence of authoritative assertions or provide for the storage of any data identifying or further qualifying authoritative assertions.

Since Kaufman fails to teach the required elements of Claim 31, reconsideration of the rejection is respectfully requested.

#### Claim 32:

As discussed above in relation to Claim 1, a mesh is a two or more dimensioned data structure. Claim 32 further defines the nature of the claimed first and second data as defining “a weighted correlation mesh interrelating the authoritative assertions of said set of authoritative assertions.”

As before, Kaufman does not teach or even recognize the existence of authoritative assertions or the use of a mesh to associate authoritative assertions. Furthermore, Kaufman does not teach the storage of such a mesh to a database.

Since Kaufman fails to teach the required elements of Claim 32, reconsideration of the rejection is respectfully requested.

#### Claim 33:

The claim requires, relative to Claim 32, that the “weighted correlations include directional information reflecting the ordered of occurrence of the authoritative assertions.” Therefore, the claim requires, as between two specific, *i.e.*, correlated, authoritative assertions, the stored data includes a specification of the direction between the two authoritative assertions.

Kaufman does not teach storing any such correlations. Kaufman does not teach any consideration of the mutual direction between any two words or sentences. In particular, none of the equations or calculations described by Kaufman include a direction term. Indeed, the word “direction” is apparently not even present anywhere within the text of the reference.

Since Kaufman fails to teach the required elements of Claim 33, reconsideration of the rejection is respectfully requested.

Claims 34 – 36:

Claim 34 further qualifies the second data as defining a correlation between “first and second predetermined authoritative assertions by a weighted ordered distance metric.”

Claim 35 requires computation of “a semantic affinity metric for the authoritative assertions.”

Claim 36 further qualifies the second data as containing “cluster association information . . . determined based on said semantic affinity metric as computed for each of the authoritative assertions within said set of authoritative assertions.”

As discussed above, Kaufman does not mention or otherwise consider direction or semantic relations in determining any correlations, let alone correlations specifically between authoritative assertions. Consequently, these claims cannot be anticipated by Kaufman. Reconsideration of the rejection of Claims 34 – 36 is respectfully requested.

Rejections under 35 U.S.C. §103:

Claims 3 – 8 and 26 – 29 stand rejected under 35 U.S.C. §103 as obvious based on the combination of Kaufman and Farahat et al (US Pub 2003/0226100).

Claims not identically shown by a reference otherwise available under 35 U.S.C. §102(a), (b), or (e) may be obvious under 35 U.S.C. §103. To establish a prima facie case of obviousness, three basic criteria must be met. First, there must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference or to combine reference teachings. Second, there must be a reasonable expectation of success. Finally, the prior art reference (or references when combined) must teach or suggest all the claim limitations. The teaching or suggestion to make the claimed combination and the reasonable expectation of success must both be found in the prior art, not in applicant’s disclosure. In re Vaeck, 947 F.2d 488, 20 USPQ2d 1438 (Fed. Cir. 1991). See also, M.P.E.P. §§2142, 2143.

Claim 26:

Claim 26 has been amended to make the claim specific to an authoritative document collection including first and second documents and where each authoritative statement includes at least one authoritative assertion and at least one authoritative citation.

Additionally, Claim 26 had been amended to further define the structure of the correlated association between authoritative assertions:

wherein said first authoritative statement includes a first authoritative citation that identifies a defined sub-portion of said second document, wherein said first authoritative assertion is semantically identified with said second authoritative assertion within said defined sub-portion of said second document

and nature of the corresponding weight values stored:

wherein said weight values include a predetermined reference weight value that interrelates a first authoritative assertion of a first authoritative statement present within said first document with a second authoritative assertion present within said second document,

and wherein said predetermined reference weight value represents the correlated semantic associativity of said first and second authoritative assertions dependent on said first authoritative citation;

Finally, the claimed viewer presents a view of a “subset” the authoritative statements and “successively selected ... supplemental” authoritative statements related to the initial subset based on the weight values stored by the second database.

Thus, of particular significance, the claim provides for the determination of a particularly defined weighting value that reflects the indirect association between authoritative assertions found to exist in different documents of the same authoritative document collection. That is, Claim 26 requires the storage of a frequency metric that

associates two assertions through a citation further refined by a found semantic association that mutually identifies the assertions.

As reviewed in detail below, Kaufman and Farahat both fail to in recognize or use any aspect distinctive of authoritative document collections. As far as Kaufman and Farahat are concerned, in terms of the operation of their systems, nothing distinguishes an authoritative document collection from any other collection. Conversely, Claim 26 specifically recognizes and utilizes the specific attributes of authoritative document collections; the claim explicitly requires an authoritative document collection.

In particular, neither Kaufman nor Farahat teaches or suggests any possible utility of associating a specific authoritative assertion that occurs in one document with an authoritative assertion that occurs in another document. While both consider and weight certain individual attributes of generic document collections, neither even identifies authoritative assertions as distinguishable, let alone having significance in relating documents.

Beyond failing to teach or suggest any way of even identifying authoritative assertions, neither Kaufman nor Farahat teaches or suggests any manner or mechanism of actually establishing a specific relation between authoritative assertions that spans documents. Kaufman only teaches direct match frequency counting within a single document at a time and Farahat only teaches regression analysis on a single document at a time against an independently formed model. Neither teach or suggest, the extended indirect association found by selective use of an associated citation and found semantic association as required by Claim 26.

The failure to teach or suggest the present invention is effectively inherent in Kaufman and Farahat, since both are merely terminal add-ons to conventional search engines. Both rely on dynamically produced search results. Both operate merely to refine the relevance ranking of the found search results. The computation results produced by the Kaufman and Farahat systems are transient and discarded with the final production of each search results list. In contrast, the weightings produced by the claimed invention are persisted to a database and provide a viewable model of the authoritative document collection.

### Kaufman

In summary, Kaufman describes a database search system where ranked search results are dynamically generated based on *ad hoc* query texts. The databases mentioned by Kaufman are described as only containing full text documents. The “authoritative” database of Kaufman is distinct from the “public” database only in the sense that the documents stored by the authoritative database are particularly selected as reliable (col 4, ll 16-18); unreliable documents are those that contain inconsistencies and errors, contain advertising, invisible words and other artifacts intended to affect word frequency, and that change (col 1, ln 48-58; col 2, ln 12-17). The authoritative database is searched in parallel with the public database as a control to “assess the relevance of the documents retrieved from the public database” (col 2, ln 31- 33).

Therefore, the content of the authoritative database is inconsequential to Kaufman provided that whatever documents are included, those documents simply act as a stable reference for evaluating the documents retrieved from the public database. Kaufman makes no distinction for documents that are part of an authoritative document collection; no aspect of the stated required “reliability” of the Kaufman databases depends on or in any way recognize authoritative statements or assertions that may exist within the documents.

The documents retrieved from the authoritative and public databases are parsed, quantized, and analyzed dynamically only after a given query text is presented. The results produced are specific to the given query text. Furthermore, given that Kaufman recognizes that the public database is continually changing, Kaufman manifests no intent or interest in to persistently retain, even temporarily, any computational results that are dependent on a query text.

### Farahat

Farahat also describes a system intended to enhance the quality of the ranking of documents returned by a search engine. For each document returned, Farahat applies a textual authority estimating model to obtain an internally-based gauge of the authority that can be assigned to the document (¶18, 40). Farahat explains this internal authoritativeness as representing whether the document appears well-researched, uses language in a “skillful and appropriate way,” and contains “numerous references of the right sort, and the like” (¶18).

The Farahat system employs a “regression or classification process” to evaluate individual documents produced by a search. The regression analysis is intended to determine the internal authoritativeness (¶159, 60) of document as a whole, one document at a time. The internal authoritativeness value determined is then used in adjusting the final ranking of the search result documents (¶162). The regression algorithms are premised on a hand-qualified set of documents that are, in turn, used to train various statistical models (¶164 – 67). These models, each considering in isolation only a single document attribute, together implement the textual authority estimating model (¶163).

As is clear from Farahat, the textual authority estimating model considers only discrete internal features of a single document at a time. In essence, the internal authoritativeness value is nothing more than a best-fit metric relating a single document against an independent, fixed document collection that is presumptively authoritative. As between two documents from a searched document collection, the Farahat system neither produces nor considers any mutually correlative metric between those particular documents; the only correlations considered are relative to an independent collection of documents.

While the use of an independent document collection against which to evaluate internal authoritativeness of individual documents is appropriate for the purposes of Farahat, the reference fails to teach or suggest any manner of correlating specific elements of two (or more) documents from the same document collection.

#### Specific Responses to the Reasons for Rejection

Contrary to the assertion made in the Action, the authoritative database of Kaufman does not correspond to either the first or second database of Claim 26. The cited section of Kaufman (col 3, ll 48-64) does not teach or suggest that authoritative statements or the results of frequency matching for a given query text are ever stored to a database. In particular, storing query results would be contrary to the stated problem faced by Kaufman; namely, that the public database changes. Storing the results would both fail to recognize new public database additions and expose the results to errors should documents in the public database be changed or removed.

Regarding the viewer element, the claim is express in establishing that the view is generated from the weights stored by the databases. As above, Kaufman actually teaches against the storing of any query dependent computations or rankings. Kaufman teaches and

suggests that the final list (view) be generated directly from a query text selected and dynamically evaluated set of documents.

Finally, Farahat does not teach or suggest anything analogous to the viewer augmenting the view presented by further selecting "supplemental authoritative statements." The section of Farahat (¶118) cited in support does mention the outputting of values. As is clear from the balance of the balance of the specification, these values are not displayed. In particular, note that corresponding step (Figure 8/ S340; ¶152-55) produces no display and the steps that do produce a display expressly produce either a reordered search list (Figure 9/S450; Figure 10/S550; Figure 12/S750) or merely suggested alternate search terms (Figure 11/S660).

Accordingly, the combination of Kaufman and Farahat, when properly considered, do not teach or suggest the present invention as set forth in Claim 26. Applicant respectfully request reconsideration of the rejection of Claim 26, as now amended.

#### Claims 27 - 29:

Claims 27 through 29 are dependent on Claim 26. For at least the reasons set forth above in regard to Claim 26, Applicant respectfully asserts that these claims are also patentable. Reconsideration of the rejection of Claims 27 through 29 is respectfully requested.

#### Claims 3 – 8:

By dependency ultimately from Claim 1, Claim 3 is specific to an authoritative document collection. As discussed above in regard to Claim 1 and Claim 26, an authoritative document collection has a defined structure that is not considered by either Kaufman or Farahat. Neither of the cited references teach or suggest that, relative specifically to an authoritative document collection, that a database store

first data identifying a set of authoritative statements ...;  
second data specifying the locations of the authoritative assertions ...; and  
third data specifying correlated associations between the authoritative  
assertions of said set of authoritative assertions.

Farahat, even to a lesser degree than Kaufman, does not teach or suggest (1) storing the identification of authoritative statements; (2) storing the location of authoritative assertions, which are part of identified authoritative statements; or (3) the generation and storage of any values representing correlated associations between the authoritative assertions.

Farahat expressly teaches reliance on the dynamic presentation of an *ad hoc* query text and performance of the corresponding search before authority estimating values can be dynamically generated. The generated authority estimating values are bulk values that reflect only an overall gauge of document authority. The generated values are used to adjust the ranking of document against document, but cannot be used to correlate any particular part of a document to a part of any other document.

Claim 1 requires storage of “correlated associations between the authoritative assertions of said set of authoritative assertions.” Therefore, “correlated associations,” where they exist, must describe associations between particular elements of the documents, *i.e.*, between particular authoritative assertions.

At a minimum, no combination of Kaufman and Farahat teaches or suggests the combination of storing the first, second and third data to a database and, further, “generate a mesh representational view of the correlated associations between the authoritative assertions.”

Claims 3 – 8, by dependence on Claim 1, are therefore patentable over Kaufman and Farahat. Applicant respectfully requests reconsideration of the rejection of Claims 3 – 8.

#### Conclusion:

In view of the above Amendments and Remarks, Applicants respectfully assert that Claims 1 – 8 and 26 – 36 are properly in condition for allowance. The Examiner is respectfully requested to take action consistent therewith and pass this application on to



issuance. The Examiner is respectfully requested to contact the Applicants' Attorney, at the telephone number provided below, in regard to any matter that the Examiner may identify that might be resolved through a teleconference with the Examiner.

Respectfully submitted,

Date: 5/16/2007

By: Gerald B. Rosenberg  
Gerald B. Rosenberg  
Reg. No. 30,320

NEWTECHLAW  
260 Sheridan Avenue, Suite 208  
Palo Alto, California 94306  
Telephone: 650.325.2100